Glossing of American Sign Language (ASL)

Author

Muhammad Arslan Thaheem

Regn. Number

317764

Supervisor

Dr. Karam Dad Kallu

MS ROBOTICS & INTELLIGENT MACHINE ENGINEERING

DEPARTMENT OF ROBOTICS & AI

SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

Nov 2022

Glossing of American Sign Language (ASL)

Author

Muhammad Arslan Thaheem

Regn Number

317764

A thesis submitted in partial fulfillment of the requirements for the degree of

MS ROBOTICS & INTELLIGENT MACHINE ENGINEERING

Thesis Supervisor:

Dr. Karam Dad Kallu

Thesis Supervisor's Signature: ____

MS ROBOTICS AND INTELLIGENT MACHINE ENGINEERING

DEPARTMENT OF ROBOTICS & AI

SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

Nov 2022

FORM TH-4

National University of Sciences & Technology

MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: (Student Name & Regn No.) Muhammad Arslan Thaheem (317764)

Titled: Glossing of American Sign Language (ASL) be accepted in partial fulfillment of the requirements for the award of MS Robotics and Intelligent Machine Engineering degree.

Examination Committee Members

1. Name: Dr. Hassan Sajid Signature:_______________

2. Name: Dr. Muhammad Jawad Khan Signature:_______________

3. Name: N/A Signature:_______________

Supervisor's name: Dr. Karam Dad Kallu Signature:_______________

Date:__________________

_____________________ _____________________

Head of Department Date

COUNTERSINGED

____

Date:__________ Dean/Principal

Thesis Acceptance Certificate

It is certified that the final copy of MS Thesis written by Muhammad Arslan Thaheem (Registration No. 317764), of Department of Robotics & AI (SMME) has been vetted by undersigned, found complete in all respects as per NUST statutes / regulations, is free from plagiarism, errors and mistakes and is accepted as a partial fulfilment for award of MS Degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in this dissertation.

Signature: ____________________

Date: ____________________

Dr. Karam Dad Kallu (Supervisor)

Signature HOD: ____________________

Date: ____________________

Signature Principal: ____________________

Date: ____________________

Declaration

I certify that this research work titled "*Glossing of American Sign Language (ASL)"* is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledged / referred.

Signature of Student

Muhammad Arslan Thaheem

317764

Plagiarism Certificate (Turnitin Report)

This thesis has been checked for Plagiarism. Turnitin report endorsed by Supervisor is attached.

Signature of Student

Muhammad Arslan Thaheem

317764

Signature of Supervisor

Dr. Karam Dad Kallu

Copyright Statement

Acknowledgements

First of all I am thankful to Allah Almighty for guiding me throughout this journey.The success of this thesis wouldn't have been possible without His help.

I would like to express my gratitude to my research supervisor, Dr. Karam Dad for his exceptional support throughout this journey. I would also like to acknowledge my committee members Dr. Hasan Sajid and Dr. Jawad Khan for their continuous help and encouragement during my journey.

I am very thankful to my parents and family members for continuously motivating and uplifting me to complete this research.

Dedication

Dedicated to my mother who would have been very happy if she was with us. After that, dedicated to my father and family for continuous support.

Abstract

American Sign Language (ASL) is a sign language used in America with slight modifications across different regions. Over the years it has developed and included a lot of new signs in it. In order for the deaf community to take notes and communicate with common people, ASL glossing is done which is an organized sentence structure of ASL. The goal of this research is to make a rule-based engine that can convert English sentences into ASL Gloss. The research included three phases. Firstly, we collected the English to ASL sentences from different resources including books, websites, etc. Secondly, we made rule-based engine to standardize the format of glossing. Since there are no proper set of rules written till now for ASL, we extracted common rules from sentences collected from different sources and continued our research on the basis of that. Thirdly, the engine was checked using different English sentences. We were able to achieve a BLEU score of 20.85 on the test set.

Table of Contents

List of Figures

List of Tables

List of Acronyms

1.	ASL	American Sign Language
2.	BSL	British Sign Language
3.	BLEU	Bilingual Evaluation Understudy
4.	ATIS	Air Travel Information System
5.	NMT	Neural Machine Translation
6.	SLT	Sign Language Translation

CHAPTER 1: INTRODUCTION

The world is changing at an extreme pace and the communication between people has entered a new era. With the advent of the internet and smartphones, people can now communicate instantly with one another over large distances however, the deaf community is still facing issues to communicate with the world even with all these technological advancements.

The deaf community still faces issues while watching videos on digital platforms as most of the videos do not have captions and even if there are captions they mostly are available in a limited number of languages. Even if the captions are available, the deaf community usually has issues comprehending them because their usual mode of communication is sign language

The deaf community usually communicates with sign language which is a visual form of communication. It is highly misunderstood that sign languages are derived from their spoken languages. Sign languages have their own grammar and linguistic features.

Just like all other languages, different sign languages are evolving day by day and from region to region. Often times there is a need to include a new sign for a new word, for example, previously there was no sign for COVID - 19 but afterward it has a specific sign in many sign languages. Also, the sign for COVID-19 varies from one sign language to another.

American Sign Language (ASL) is a sign language signed in most of America. It is surprising to know that even ASL has different variations in different regions of America. ASL is being taught in different institutes and organizations but as mentioned above communication has always remained a challenge for the deaf community, especially when it comes to the communication between the deaf community and common people. To fill this communication gap and to make communication faster we have made a rule-based engine that can easily convert an English sentence to its gloss.

Our rule-based engine is a part of a project where the English sentence is first converted to its gloss and then to ASL. Conversion to gloss is necessary as gloss helps to remove the unnecessary words and do any other changes needed for the sign language.

1.1 Problem Statement

Traditionally ASL is being taught in a conventional way in different educational institutes for deaf people. Since ASL resources are limited only to the people who are connected to this field like ASL teachers and students, the communication between common and deaf people is always a big challenge. To our knowledge, there is no current platform that supports the automated glossing of English text which can lead to a new era of communication within the deaf community.

1.2 Proposed Solution

As a solution we propose to make a rule-based engine that can automatically convert an English sentence into ASL gloss. Although there are apps, and websites, which teach ASL, just like any other language learning ASL can also take time but our rule-based engine will provide beginners and common people to communicate effectively with the deaf community. Since the application will give ASL gloss on the run time so this will effectively reduce the time that might be wasted otherwise if a common and a deaf person are trying to communicate.

1.3 Expected Outcome

The aim of this project is to solve the problem of the communication gap between common people and the deaf community. The first outcome of the research is expected to compile real English to gloss examples from different sources like websites, books, etc. The second expected outcome is the set of standardized rules that can later be used as a reference for further research. The third outcome is to make a rule-based engine that can automatically convert text to gloss and facilitate communication between common people and the deaf community.

1.4 Methodology

The research was conducted in three phases. Firstly, we had to manually collect English to ASL sentences from different books, workbooks, websites, etc. We collected data of 837 examples from almost 5 different resources. Secondly, the available data was then divided into 3 types namely wh-questions, polar questions, and simple sentences. Thirdly, linguistic features were then extracted from the above sentence types using the spaCy library. Then, by comparing the English text and its respective gloss, the linguistic features like parts of speech, tags and dependencies, etc were used to specify the rules.

1.5 Thesis Overview

This thesis is divided into the following chapters, firstly the thorough study of literature available publicly is presented in detail in the **Literature Review** chapter. The chapter also highlights the problems associated with the machine learning approach to the problem. Afterward, in the **Methodology** chapter, a thorough analysis of the problem at hand is mentioned. The chapter also includes the details of gathering the dataset and the technical problems associated with that. The last part of the chapter explains the making of the rules according to the linguistic features extracted from the spaCy library. The outcomes from different types of English sentences are presented in the **Results** chapter, which also includes the BLEU score of the rule-based engine. **Future Work** chapter discussed the research ideas that can be implemented after this research. Finally, the **Conclusion** chapter discusses the overall research and it’s outcomes.

CHAPTER 2: LITERATURE REVIEW

We have divided the Literature review into the following three parts:

1. Related Work
2. Discussion on Datasets
3. ASL Rules

2.1 Related Work

Communication by a system of gestures is not exclusively human activity so in a broad sense of the term, sign language is as old as the race itself and its earliest history is equally obscure [1].

Stokoe [1] proposed the first annotation system for writing American Sign Language using graphical symbols.

Tmar at al. [2] addressed the problem of the lack of a large parallel corpus in the field of American Sign Language translation. They used a rule-based approach to convert a large English text into an ASL gloss dataset called the ASLG-PC12 dataset. Before this paper, a lot of research on ASL was based on video recordings. They made the dataset public thus paving way for advanced search in the field of ASL. A lot of research work mentioned below used the ASLG-PC12 dataset, as a base dataset for their paper.

Bungeroth at al. [3] presented the ATIS sign language corpus for five languages, English, German, Irish sign language, German Sign Language, and South African Sign Language. The corpus contains English phrases and sentences related to travel information as the corpus is based on the Air Travel Information System (ATIS) dataset.

Bonham [4] used the small parallel corpus of English text and American Sign Language gloss from The Church of Jesus Christ of Latter-day Saints. After cleaning the corpus, they trained the corpus on a Moses MT system. After fine-tuning and several iterations, they were able to yield a good BLEU score.

Manzano [5] made a video-to-video translation system that would generate a puppet interpreter for a given video and translate the speech signals into ASL. The main steps were to get an audio transcription of the video and then use a Neural Machine Translation (NMT) module to convert text to ASL. Lastly, an avatar would make the signs. The data used was

ASLG-PC12 and the BLEU score attained was 17.73.

Arvanitis at al. [6] used a sequence-to-sequence attention mechanism to convert ASL gloss to text. Using the attention mechanism helped to align the encoder and decoder hidden states. They implemented a sequence-to-sequence attention system, using two different architectures with promising results. The data used was ASLG-PC12 and the BLEU score attained was 0.65

Kayahan at al. [7] proposed a hybrid translation approach to translating the Turkish Language into Turkish Sign Language (TID). Turkish text is first translated from Turkish to Turkish Sign Language gloss using pre-defined rules. After that, the translated text is fed into a statistical translation component to complete the translation process. They used an online Turkish to TID dictionary built by the Turkey Ministry of Family and Social Policies. The dictionary contained video and gloss representations of the TID signs. It also had Turkish to TID sample sentences with relevant glosses. They created a parallel corpus using sample sentences from each word translation. The task was accomplished by combining 2000 alphabetically grouped words and a website crawler to extract 3561 sentence pairs.They defined 13 rules to translate Turkish to TID. They were able to achieve a BLEU score of 12.64 using this hybrid approach.

San-Segundo et al. [8] proposed a translation system for deaf people. They implemented a brilliant idea of helping the deaf community by making a system that would help the deaf community for making or renewing ID cards. They focused on data containing sentences spoken by an employee while helping people for making ID cards.The translation was divided into 3 parts. Firstly, a speech recognizer would convert speech to text/words. Secondly, the text was converted to sign language by implementing rule-based and statistical methods separately, using a natural language translator. Lastly, a 3D avatar would convert the sign language to signs in 3D. The system was able to achieve a BLEU score of 0.578 using a rule-based approach while a score of 0.4941 using the statistical translator approach. The paper also addresses the delay issues between the spoken utterance and the 3D animation.

Stoll et al. [9] proposed a novel approach to using Generative Adversarial Networks (GANs) to create their own sign language video instead of relying on the traditional approach of using avatars for showing sign language. This approach made it possible to achieve the task by using minimum annotated data and gloss. They divided the task into two parts. Firstly, the text was converted to gloss using Natural Language Translation component. Secondly, after the data-driven mapping, they used GANs to make sign language video sequences. For spoken-to-sign language gloss, they employed encoder-decoder architecture [25] with Luong's attention [26]. The data used by them was the PHOENIX-Weather 2014T Sign Language Translation dataset reporting a BLEU-4 score of 16.34/15.26 on dev/test sets. Due to the limited number of signers in the PHOENIX-Weather 2014T dataset, they used another large dataset, the SMILE Sign Language Assessment dataset [27], to train the multi-signer (MS) generation network.

Kayo Yin et al. [10] introduced the STMC-Transformers obtaining the BLEU score of 5 and 7 using gloss-to-text and video-to-text translation on the PHOENIX-Weather 2014T dataset. Also, they obtained a BLEU score of over 16 on the ASLG-PC12 dataset. They have also claimed that weight-tying or pre-trained embeddings GloVe3 [28] and fastText [29] that are used by them on PHOENIX-Weather 2014T have never been employed in Sign Language Translation SLT.

Jiangbin Zheng et al. [11] identified that SLT systems based on neural translation frameworks recently used have made progress but they do not perform as efficiently on long sentences that often require long-distance dependencies.They tackled the problem in two steps, Firstly they used a frame stream density compression (FSDC) for shortening long sign sentences without losing information. Secondly, the traditional encoder is improved by using a temporal convolution (T-Conv) unit and a dynamic hierarchical bidirectional GRU (DH-BiGRU) unit sequentially. The dataset used was PHOENIX-Weather 2014T data and they were able to attain a 1.5+ BLEU score gain as compared to state-of-the-art models.

2.2 Discussion on Datasets

A parallel corpus is a large and structured text aligned between source and target languages.

They are used to do statistical analysis and hypothesis testing, checking occurrences, or validating linguistic rules on a specific universe. [2]

In Neural Machine Learning, the availability of a large parallel corpus is necessary. While for spoken languages large amounts of the parallel corpus are available online for Neural Machine Translation, such is not the case for Sign Languages.

Manzano [5] pointed out that there is no proper dataset for sign language translation. The existing ones are very small that can affect the training process in NMT.

In the past, several attempts have been made to convert video recordings into ASL gloss. According to [6], this process is done in the following steps:

1. Capture video from signers
2. Extract features from processed frames

3. Map features to corresponding text

Few of the datasets used in the field of Sign Language are described below:

1. Tmar at al. [2] used a rule-based approach to covert a large English text into an ASL gloss dataset called the ASLG-PC12 dataset. ASLG-PC12 contains 87710 parallel sentences. Since 2013, this dataset is widely used in the field of American Sign Language.
2. Bungeroth at al. [3] presented a sign language corpus of five languages, called ATIS Air Travel Information System (ATIS) dataset. The corpus is mostly related to Air travel Information.
3. Forster et al. [12] introduced the PHOENIX-Weather 2014 dataset. Their strategy included gathering data from German public TV and weather forecasts of a subset of 386 editions were transcribed using gloss notation. Using an open-source speech recognizer they transcribed spoken German from videos.
4. Sang-Ki Ko et al. [13] created their own video dataset called the KETI (Korea Electronics Technology Institute) Sign Language dataset, which consists of 14,672 videos of high resolution and quality.
5. Dilek Kayahan at al. [7] composed the Turkish to TID parallel corpus using an online Turkish to TID dictionary built by the Turkey Ministry of Family and Social Policies.
6. Bonham [4] mentions a Text-to-Gloss dataset from The Church of Jesus Christ of Latter-day Saints.
7. The European Cultural Heritage Online organization (ECHO) published corpora for 3 sign languages namely British, Swedish, and Netherlands. [14,15,16]
8. An ASL Linguistic Research group from Boston University published a corpus in American Sign Language [17].

2.3 ASL Rules

It is important to understand the fact that a spoken language and its sign language are grammatically not the same. Also, the sign languages vary from area to area for example the American Sign Language (ASL) is different from British Sign Language (BSL).

Upon starting this project, the first and foremost hurdle was to find the standards for American Sign Language (ASL). According to [5] the absence of a global standard in sign language makes sign language-related tasks very difficult.

[18] has excellently demonstrated the issue of the grammatical difference between an English text and its gloss using the figure below

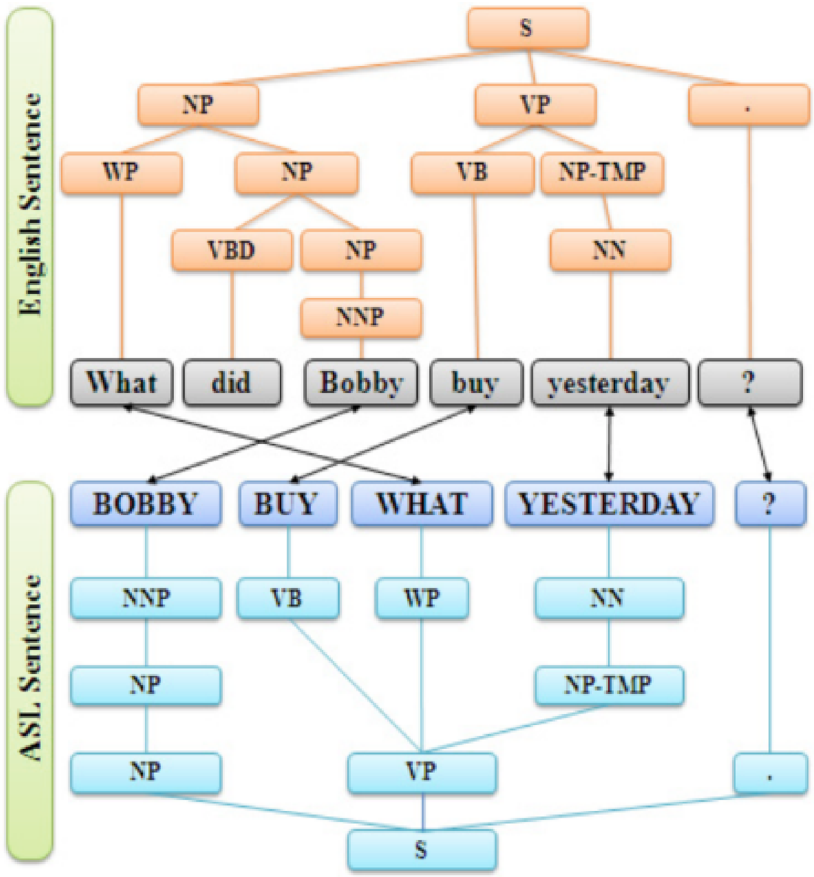

Figure 2.1: Grammatical difference between ASL gloss and English sentence

In an effort to find standards/rules for ASL we managed to contact some professionals in America who are working in institutions related to the deaf community or have worked in the field of ASL. We got some very good answers from those individuals. We will be adding those responses here

2.3.1 First Response

“Arslan

I am glad to see someone like you caring about “defining ASL”. I am not an ASL specialist or educator. I am a retired librarian and a library advocate Supporting quality deaf cultural books and programs in local communities of the nation, the public library. My author friend Nancy has been working on a book about Laurent CLERC who invented ASL. We have been searching for how we should define ASL. She uses the Oxford University Dictionary. And she also added a true fact about it.   Let me know if you are interested in helping Friends of Library for Deaf Action with the best definition. ”

2.3.2 Second Response

“Alice,

Just a reminder:
No one person ever "invents" a language. Laurent Clerc did not "invent" ASL.

His French sign language influenced the development of American Sign Language, along with the deaf students' "home signing" that they brought from their homes. All languages develop naturally over time among groups of people.”

2.3.3 3rd Response

“Arslan

Yours is a good question that gets asked rather often. Chris made an important point and I would like to add a bit to that as a deaf observer and advocate, and not as an ASL or language expert. Language development is a fascinating thing and I wish I knew more but the essence of what Chris is saying is that the development is much like that of regional accents or dialects; different areas have different signs for the same thing. Signs can be disparate even within a given state or region and there are naturally a host of other factors that make ASL a very dynamic language that is continually evolving.

The advent of social media is stirring that "melting pot" we have and we are creating certain signs that get invented on the fly, like Covid-19 (an open hand rotating behind a closed fist to mimic the spike proteins), and these often become adopted signs nationally, even internationally. I believe this mimics regular language use or the lingua franca whereas certain words enter a common lexicon or common use that is then sprinkled into whatever dialect is being used. In other words, you'd recognize some signs but be flummoxed by others.

Naturally, we defer to organizations like ASLTA, ASL Teachers Association, RID, Registry of Interpreters of the Deaf, NAD, or National Association of the Deaf for language use and norms but standardization has always been a challenge. Personally, I think that interpreter training programs should have at least two distinct parts --one national and one regional-- because they effectively deal with different languages and language structures; there is one type you need in order to interpret on a national phone relay company (VRS or Video Relay Services) and another type you need in order to provide regional in-person services in Murfreesboro, Tennessee. This means that if you learn in one area or state and then move to another, you will want to take some time to acclimate to the regional ASL dialect --use expressions, idioms, and the like.

Finally, the dearth of information is precisely what Alice is working on when she talks about NDHM, National Deaf History Month, DCDL, Deaf Culture Digital Library, library programming, or just any old library. Anything you find in your studies can potentially contribute to content development, e.g., books, studies, and/or stories that further common understanding and/or sharing. I am sure people here would be thrilled to see what discoveries you might make in pursuit of ASL and the challenges posed by the concept of language"

From the above conversation and self-study, we concluded that there is no proper set of standards for ASL gloss and that we will have to consult different sources like books, websites, etc for extracting common rules. The next chapter explains the methodology to come up with a set of rules in ASL.

CHAPTER 3: METHODOLOGY

3.1 Data Collection

As mentioned in the sub-section of Chapter 2, “Discussion on Datasets”, it has always been challenging for researchers to have a good dataset to conduct research in the field of ASL.

The two most common datasets that have been used in the research are the ASLG-PC12 dataset and PHOENIX-Weather 2014 dataset. These two data sets are very important in the field of ASL. This can be seen from the fact that the above two datasets are used, in a total of 6 times in the research work mentioned in the sub-section of Chapter 2, “Related Work”.

Before we start explaining our data collection process it is important to understand the need for custom data collection and why we didn’t use the already present datasets. This project is a sub-project of a big project that aimed to minimize the communication gap between the deaf community and common people. The deliverable of the project was to make a system that can convert an English text into gloss and later the gloss can be converted into signs.

As a part of this big project, we were assigned the task of Text-to-Gloss. After the literature review, we used the common approach used by other researchers. We used the ASLG-PC12 dataset and applied Neural Machine Translation using the transformers-based approach with an attention mechanism [30]. The accuracy results that we achieved were quite promising but there were a few limitations mentioned below

1. ASLG-PC12 dataset was created using a rule-based algorithm so the accuracy of the Neural Machine Translator was quite promising but the data set obviously doesn't include all the vocabulary of the English language. The requirement of our application was to convert any English text to gloss but the limited vocabulary of the data posed a limitation.
2. The rules of ASL and the grammatical structure were not quite clear that actually limits a reader to have a full grasp of the subject.

Some of the results of Neural Machine Translator are given in Table 3.1

Example 1	Input	together with the other countries of europe , we will be able to overcome the challenges we currently face .
	Ground Truth	desc-toger with desc-or country europe , x-we will be desc-able to overcome challenge x-we desc-currently face .
	Prediction	desc-toger with desc-or country europe, x-we will be desc-able to overcome
Example 2	Input	ask the student where he lives.
	Ground Truth	ask-him student where ix-he live
	Prediction	ask student where x-he life.

Table 3.1: Results of Neural Machine Translator

The above-mentioned reasons compelled us to tackle the given problem with a different approach than Neural Machine Translation. This led us to move toward the rule-based approach that was originally used to create the ASLG-PC12 dataset. The first task in making a rue-based engine was to collect real examples of English to gloss. As mentioned in the sub-section of Chapter 2, "ASL Rules", the ASL is evolving day by day, and we would need to consult multiple sources to extract data and ASL rules. Therefore, upon searching we found some very good resources from which we manually extracted text-to-gloss data. The sources are mentioned below

1. Life Print Website (https://www.lifeprint.com/)
2. The American Sign Language Phrase Book, Third Edition, by Barbara Bernstein Fant, Betty Miller, Lou Fant.
3. ASL Grammar - The Workbook (2018), by Rochelle Barlow
4. ASL Workbook by Grace Chapel Deaf Ministry
5. Signing Savvy Website (www.signingsavvy.com)

Collecting the data manually from different sources was a cumbersome task as not all the books available on the subject had proper text-to-gloss examples. Out of all the books, websites, etc., we were able to extract relevant examples from the above-mentioned 5 sources. By the end, we were able to collect 837 examples. (587 samples were used to create the rules and 250 samples were used for testing)

It is important to mention here that for a machine learning task the number of extracted examples won't be enough but for a rule-based approach, these 837 examples proved to be sufficient for extracting many rules of ASL glossing.

3.2 Sentence Division

After collecting the data and thoroughly analyzing the data, we divided the data into 3 main types of sentences namely:

1. Wh - Questions
2. Polar Questions
3. Simple Sentences

The sentences are further divided into their subcategories in order to divide the problem into subparts for easy dealing with sentences and for making different rules.

It is important to mention here that the division is based purely on the collected data. Although it covers most variations of English sentences, adding all types of variations is quite impossible. However, the rule-based engine is made in a way that it will cater to almost any kind of single English sentence. The different variations of sentences are listed in the table below

1st Category	2nd Category	3rd Category	English Sentence
Wh - Questions	Regular Wh - Questions		What kind of soup do you like?
	How Wh - Questions		How often do you go to the library?
	Choice Wh - Questions		Which do you prefer to drink, water, milk, pop, or beer?
Polar Questions	Regular Polar Question		Do you have a backpack?
	If Polar Question	Regular If Question	If your dog gets sick do you take it to the veterinarian?
		Choice If Question	If you go to church, do you wear pants or do you wear a dress?
	Choice Polar Question		Do you prefer hamburgers or hotdogs?
	And Polar Question		Do you like green eggs and ham?
Simple Sentences	Regular Sentence		Yesterday I bought a dog.
	Rhetorical Sentence	Regular Rhetorical Sentence	My sister got into a car accident because of black ice.
		And Rhetorical Sentence	I love the fall because of the rain and the wind.
	But Sentence		His wife has red curly hair, but I don't know her name.
	Comma Separated Sentence		I want a new bowl, this one is old.
	And Sentence	Type 1	The moon and stars are bright tonight.
		Type 2	The mountain is here and the farm is just below it.
		Type 3	He pays me the money and I go out and buy the food.

Table 3.2: Variations of English sentences in our data

3.3 Preprocessing

In the processing part the data was analyzed properly and we found the following two bottlenecks

1. Contractions
2. Multiple sentences/questions in one example

3. Fingerspelling and body movement information

3.3.1 Contractions

Contractions are the shortened form of words. The collected data had a lot of contractions like don't, didn't, etc. that were removed using the "contractions" library of python. A few examples of removed contractions are given below

Examples from data	Contractions Removed
I don't care about that class.	I do not care about that class.
I don't mind it's better than being late.	I do not mind it is better than being late.
Why didn't you tell me?	Why did not you tell me?
I didn't tell him.	I did not tell him.

Table 3.3: Examples of Contractions

3.3.2 Multiple sentences/questions in one example

Although most of the data contained single sentences in one example but few examples had more than one sentence. This posed a limitation as we designed our rule-based engine to process one sentence at a time to give better results. Hence, the examples with more than one sentence were omitted or reduced to one sentence.

Some of the examples with more than one sentence are given below

Examples from data
Do you have a car?--if so--How many doors?
Do you have a pet? What is its name?
I'm going to a party Saturday night. Do you want to go with me?
if so) WHAT-huh? What, if any, habits do you have?

Table 3.4: Examples of Multiple sentences/questions in one example

3.3.3 Fingerspelling and body movement information

We removed the finger spelling and body movement information from the human-translated gloss as it was not needed for our project.

Original Gloss	After Preprocessing
SUMMER [bodyshift]-OR WINTER, YOU LIKE BETTER WHICH?	SUMMER OR WINTER, YOU LIKE BETTER WHICH?
YOUR HAMBURGER WANT M-A-Y-O YOU?	YOUR HAMBURGER WANT MAYO YOU?

Table 3.5: Gloss preprocessed for fingerspelling and body movement information

3.4 Linguistic Features using spaCy Library

As the primary purpose of this project is to make a rule-based engine for converting English text into ASL Gloss, so in our search for making such an engine, we looked for different Natural Language Processing NLP libraries in python. Some of the amazing libraries in the field of NLP are

1. NLTK
2. Stanford CoreNLP
3. spaCy
4. Textblob

Out of the above libraries we selected spaCy for making our rule-based engine. The main reason for selecting spaCy is that it is much faster and provides a range of linguistic features that can be used for making an efficient rule-based engine.

The main functionalities of spaCy that were used in our rule engine are as follows:

1. English Model - Small
2. Parts-of-speech tagging

3.4.1 English Model - Small

spaCy already has pre-trained models for English. We have used the small English model namely “en_core_web_sm”. The reason for using this model is that it works perfectly on the CPU and it is lightweight. The figure from [19] shows the details of the model

en_core_web_sm

RELEASE DETAILS

Latest: 3.4.0

English pipeline optimized for CPU. Components: tok2vec, tagger, parser, senter, ner, attribute_ruler, lemmatizer.

LANGUAGE	EN English
TYPE	CORE Vocabulary, syntax, entities
GENRE	WEB written text (blogs, news, comments)
SIZE	SM 12 MB
COMPONENTS ?	tok2vec, tagger, parser, senter, attribute_ruler, lemmatizer, ner
PIPELINE ?	tok2vec, tagger, parser, attribute_ruler, lemmatizer, ner
VECTORS ?	0 keys, 0 unique vectors (0 dimensions)
SOURCES ?	OntoNotes 5 (Ralph Weischedel, Martha Palmer, Mitchell Marcus, Eduard Hovy, Sameer Pradhan, Lance Ramshaw, Nianwen Xue, Ann Taylor, Jeff Kaufman, Michelle Franchini, Mohammed El-Bachouti, Robert Belvin, Ann Houston) ClearNLP Constituent-to-Dependency Conversion </> (Emory University) WordNet 3.0 (Princeton University)
AUTHOR	Explosion
LICENSE	MIT

Figure 3.1: Details of small English model of spaCy

3.4.2 Parts-of-speech tagging

[20] explains that after tokenization, spaCy can parse and tag a given Doc. Predictions are made using a trained pipeline and its statistical models. The predictions tell us about the tags or labels present in the text. A trained component includes binary data that is produced by showing a system enough examples for it to make predictions that generalize across the language – for example, a word following "the" in English is most likely a noun. Figure from [20] below shows the predictions on a text "Apple is looking at buying U.K. startup for $1 billion"

TEXT	LEMMA	POS	TAG	DEP	SHAPE	ALPHA	STOP
Apple	apple	PROPN	NNP	nsubj	Xxxxx	True	False
is	be	AUX	VBZ	aux	xx	True	True
looking	look	VERB	VBG	ROOT	xxxx	True	False
at	at	ADP	IN	prep	xx	True	True
buying	buy	VERB	VBG	pcomp	xxxx	True	False
U.K.	u.k.	PROPN	NNP	compound	X.X.	False	False
startup	startup	NOUN	NN	dobj	xxxx	True	False
for	for	ADP	IN	prep	xxx	True	True
$	$	SYM	$	quantmod	$	False	False
1	1	NUM	CD	compound	d	False	False
billion	billion	NUM	CD	pobj	xxxx	True	False

Figure 3.2: An example of linguistic features predicted by spaCy

The different terminologies in the header of the above table are explained below. According to [20]

Text: The original word text.
Lemma: The base form of the word.
POS: The simple UPOS part-of-speech tag.
Tag: The detailed part-of-speech tag.
Dep: Syntactic dependency, i.e. the relation between tokens.
Shape: The word shape – capitalization, punctuation, digits.
is alpha: Is the token an alpha character?
is stop: Is the token part of a stop list, i.e. the most common words of the language?

Figure 3.3: spaCy terminologies

For further understanding, we used the displaCy visualizer that uses spaCy to display the POS and dependencies in a sentence. In the figure below the predictions below the sentence are POS and above the sentence are dependencies

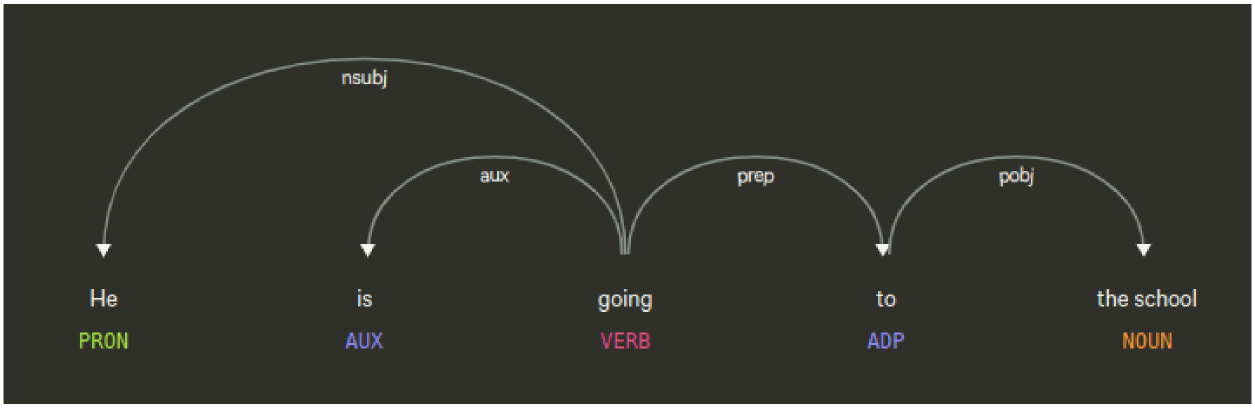

Figure 3.4: Visual representation of linguistic features by spaCy

[21,22] give the idea of labels that we used to make our rule-based engine. To the best of our knowledge, POS, tags, and dependencies combine to make 110+ labels that helped us, in different combinations, to mkae our rule engine.

POS	DESCRIPTION	EXAMPLES
ADJ	adjective	*big, old, green, incomprehensible, first*
ADP	adposition	*in, to, during*
ADV	adverb	*very, tomorrow, down, where, there*
AUX	auxiliary	*is, has (done), will (do), should (do)*
CONJ	conjunction	*and, or, but*
CCONJ	coordinating conjunction	*and, or, but*
DET	determiner	*a, an, the*
INTJ	interjection	*psst, ouch, bravo, hello*
NOUN	noun	*girl, cat, tree, air, beauty*
NUM	numeral	*1, 2017, one, seventy-seven, IV, MMXIV*
PART	particle	*'s, not,*
PRON	pronoun	*I, you, he, she, myself, themselves, somebody*
PROPN	proper noun	*Mary, John, London, NATO, HBO*
PUNCT	punctuation	*., (,), ?*
SCONJ	subordinating conjunction	*if, while, that*
SYM	symbol	*$, %, §, ©, +, −, ×, ÷, =, :), 😝*
VERB	verb	*run, runs, running, eat, ate, eating*
X	other	*sfpksdpsxmsa*
SPACE	space	

Figure 3.5: Different Parts of Speech labels by spaCy

TAGGER	`$` , `''` , `,` , `-LRB-` , `-RRB-` , `.` , `:` , `ADD` , `AFX` , `CC` , `CD` , `DT` , `EX` , `FW` , `HYPH` , `IN` , `JJ` , `JJR` , `JJS` , `LS` , `MD` , `NFP` , `NN` , `NNP` , `NNPS` , `NNS` , `PDT` , `POS` , `PRP` , `PRP$` , `RB` , `RBR` , `RBS` , `RP` , `SYM` , `TO` , `UH` , `VB` , `VBD` , `VBG` , `VBN` , `VBP` , `VBZ` , `WDT` , `WP` , `WP$` , `WRB` , `XX` , `_SP` , ` `` `
PARSER	`ROOT` , `acl` , `acomp` , `advcl` , `advmod` , `agent` , `amod` , `appos` , `attr` , `aux` , `auxpass` , `case` , `cc` , `ccomp` , `compound` , `conj` , `csubj` , `csubjpass` , `dative` , `dep` , `det` , `dobj` , `expl` , `intj` , `mark` , `meta` , `neg` , `nmod` , `npadvmod` , `nsubj` , `nsubjpass` , `nummod` , `oprd` , `parataxis` , `pcomp` , `pobj` , `poss` , `preconj` , `predet` , `prep` , `prt` , `punct` , `quantmod` , `relcl` , `xcomp`

Figure 3.6: Different tags and dependencies by spaCy

3.5 ASL Rules

As mentioned in the sub-section of Chapter 2, "ASL Rules" and sub-section of Chapter 3, "Data Collection", ASL is a dynamic language and there is no proper set of rules/ standard for ASL gloss. ASL also varies from area to area and it is continuously evolving. Also, the datasets that exist for ASL do not incorporate/explain all the rules in detail.

In an effort to find a standard for ASL glossing, we found the workbook "ASL Grammar - The Workbook" (2018), by Rochelle Barlow [23]. Rochelle has actually done a great job of bringing ASL into a structured form.

The main rules given by Rochelle [23] are given

TIME+ TOPIC+ COMMENT+ QUESTION

TIME+ TOPIC+ COMMENT+ REFERENT+ QUESTION

TIME+ TOPIC+ COMMENT+ QUESTION + REFERENT

TIME+ TOPIC+ COMMENT+ ACTION +QUESTION

TIME+ TOPIC+ COMMENT+ QUESTION +COMMENT

An example is given below for a further understanding of the rules

Can you teach me sign language?

Time = none
Topic = sign language
Referent = you, me
Comment = teach
Question = can you

We take the following points from the above rules

1. The rules are very well designed but after looking at our data we came to the idea that they are a bit generalized version of rules.
2. An example is that the “COMMENT” includes an adjective, adverb, or verb but there are a lot of different English sentences with different types of grammatical labels that is why a detailed version of the rules is required.
3. The rules do not cover, in detail, some complex English sentences mentioned in Table 3.2.

Because of the above points and keeping into consideration the different types of sentences mentioned in Table 3.2, we concluded that detailed standards for ASL glossing are needed.

For making a rule-based engine with ASL standards we used the data from 5 sources mentioned in the sub-section of Chapter 3, “Data Collection”. The approach was to divide the sentences into 1st, 2nd, and 3rd categories mentioned in Table 3.1 and then set standards, according to the given human-translated glossing as a reference, using the labels used in spaCy library as mentioned in the sub-section of Chapter 3, “Linguistic Features using the spaCy library”.

The rules that we have set are basically divided into two major categories

a. Wh - questions
b. Polar questions and simple sentences

We have used our own label names to define the rules. Also, some of the labels are named the same as predicted by spaCy English model. It is important to note that, we are mentioning the main labels that are important as mentioning all the rules and labels won’t be possible here.

3.5.1 Wh - Questions Rules

The rules for wh - questions are given below

TIME_PREFIX + TIME + [POINT_AT + POINT_AT_OBJECT + POINT_AT_ADJECTIVE] + [POSSESSIVE + POSSESSIVE_HEAD + POSSESSIVE_ADJ + POSSESSIVE_OBJECT] + DIRECT_OBJECT + NMOD + [COMPOUND + COMPOUND_ATTRIBUTE / NOUN] + [CLOCK_PREFIX + CLOCK] + OBJECT_OF_PREPOSITION + NOMINAL_SUBJECT + NEGATIVE + MAIN_VERB + XCOMP + ADVERBIAL_MODIFIER + ADJECTIVAL_COMPLEMENT + DATIVE + [QUESTION + HOW_QUESTION_LABEL+ QUESTION_ATTRIBUTE]

Few examples are given below.

Sentence	Gloss	Gloss divided into words	Labels
Where is my phone?	my phone where?	my	POSSESSIVE
		phone	POSSESSIVE_HEAD
		where?	QUESTION
What did you do in school?	school you do what?	school	OBJECT_OF_PREPOSITION
		you	NOMINAL_SUBJECT
		do	MAIN_VERB
		what?	QUESTION

Table 3.6: Examples of rules for wh - questions

3.5.2 Polar Questions and Simple Sentences Rules

The rules for Polar Questions and Simple Sentences are given below

TIME_ADVERBIAL_MODIFIER + TIME_PREFIX + TIME + [POINT_AT + POINT_AT_OBJECT + POINT_AT_ADJECTIVE] + [POSSESSIVE + POSSESSIVE_ADJ + POSSESSIVE_OBJECT] + APPOSITIONAL_MODIFIER + NUMERIC_MODIFIER + DIRECT_OBJECT + NMOD + [COMPOUND + COMPOUND_ATTRIBUTE / NOUN] + [CLOCK_PREFIX + CLOCK] + NOMINAL_SUBJECT + NEGATIVE + MAIN_VERB + XCOMP + ADVERBIAL_MODIFIER + ADJECTIVAL_COMPLEMENT + [ADJECTIVE + OBJECT_OF_PREPOSITION] + DATIVE + [QUESTION + HOW_QUESTION_LABEL+ QUESTION_ATTRIBUTE]

Few examples are given below.

Sentence	Gloss	Gloss divided into words	Labels
The moon is beautiful tonight.	tonight moon beautiful.	tonight	TIME
		moon	NOMINAL_SUBJECT
		beautiful	ADJECTIVAL_COMPLEMENT
Do you like to cook?	you like cook?	you	NOMINAL_SUBJECT
		like	MAIN_VERB
		cook?	XCOMP

Table 3.7: Examples of rules for polar questions and simple sentences

CHAPTER 4: RESULTS & DISCUSSION

4.1 Performance Metrics

There are two major methods to evaluate a translation task. It can either be done by a human or by using the performance metrics scores for machine translation. One of the most used scores is the BLEU score.

According to [24] BLEU (Bilingual Evaluation Understudy) score, indicates similarity between two texts, with values closer to one representing more similar texts. BLEU score gives an overall evaluation of Machine Translated text.

BLEU score is measured on a scale of 0 to 1, zero means that there is no similarity between the two texts (low quality) and 1 means the two texts are completely similar (high quality) .

The mathematical representation of the BLEU score is given below

Mathematically, the BLEU score is defined as:

$$\text{BLEU} = \underbrace{\min\Big(1, \exp\big(1 - \frac{\text{reference-length}}{\text{output-length}}\big)\Big)}_{\text{brevity penalty}} \underbrace{\Big(\prod_{i=1}^{4} precision_i\Big)^{1/4}}_{\text{n-gram overlap}}$$

with

$$precision_i = \frac{\sum_{\text{snt}\in\text{Cand-Corpus}} \sum_{i\in\text{snt}} \min(m^i_{cand}, m^i_{ref})}{w^i_t = \sum_{\text{snt'}\in\text{Cand-Corpus}} \sum_{i'\in\text{snt'}} m^{i'}_{cand}}$$

where

- m^i_{cand} is the count of i-gram in candidate matching the reference translation
- m^i_{ref} is the count of i-gram in the reference translation
- w^i_t is the total number of i-grams in candidate translation

The formula consists of two parts: the brevity penalty and the n-gram overlap.

Figure 4.1: BLEU Score Formula

4.2 Performance Metric Results

The BLEU score was calculated on test data. The user has to add the human-translated text and machine-translated text as separate files and then the user can get the result immediately.

The BLEU score that we achieved is given below

BLEU Metric	Score
BLEU-1	68.12
BLEU-2	44.22
BLEU-3	29.76
BLEU-4	20.85

Table 4.1: BLEU scores for our test data

[7] has compared its BLEU score results with [8], [5] and [9]. Just like [7], for our understanding, we will be naming different contributions as System-1, System-2, etc. Also, we added some more papers for detailed comparison. The naming convention is given below

System	Research	Sign Language
System-1	San-Segundo et al. [8]	Spanish
System-2	Our Rule-Based Engine	American
System-3	Manzano [5]	American
System-4	Bonham [4]	American
System-5	Stoll et al. [9]	German
System-6	Kayahan at al. [7]	Turkish

Table 4.2: Naming convention for different researches

The results are given below (Systems 1, 3 and 4 only have BLEU-4 score)

System	BLEU-1	BLEU-2	BLEU-3	BLEU-4
System-1 (San-Segundo et al. [8])	0	0	0	57.8
System-2 (Our Rule-Based Engine)	68.12	44.22	29.76	20.85
System-3 (Manzano [5])	0	0	0	17.73
System-4 (Bonham [4])	0	0	0	17.65
System-5 (Stoll et al. [9])	50.67	32.25	21.54	15.26
System-6 (Kayahan at al. [7])	53.17	31.48	19.28	12.64

Table 4.3: Comparison of our BLEU scores with related studies

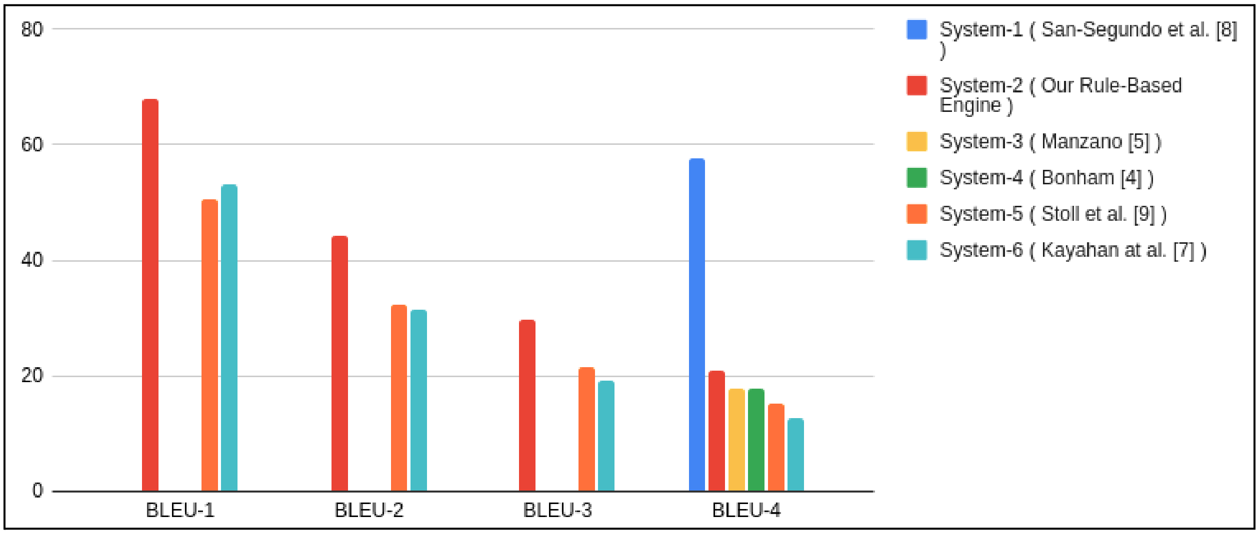

Figure 4.2: Comparison of our rule-based engine with related studies

The comparison of ASL based studies is given below

System	BLEU-4 Score	Sign Language
System-2 (Our Rule-Based Engine)	20.85	American
System-3 (Manzano [5])	17.73	American
System-4 (Bonham [4])	17.65	American

Table 4.4: Comparison of our BLEU scores with ASL based studies

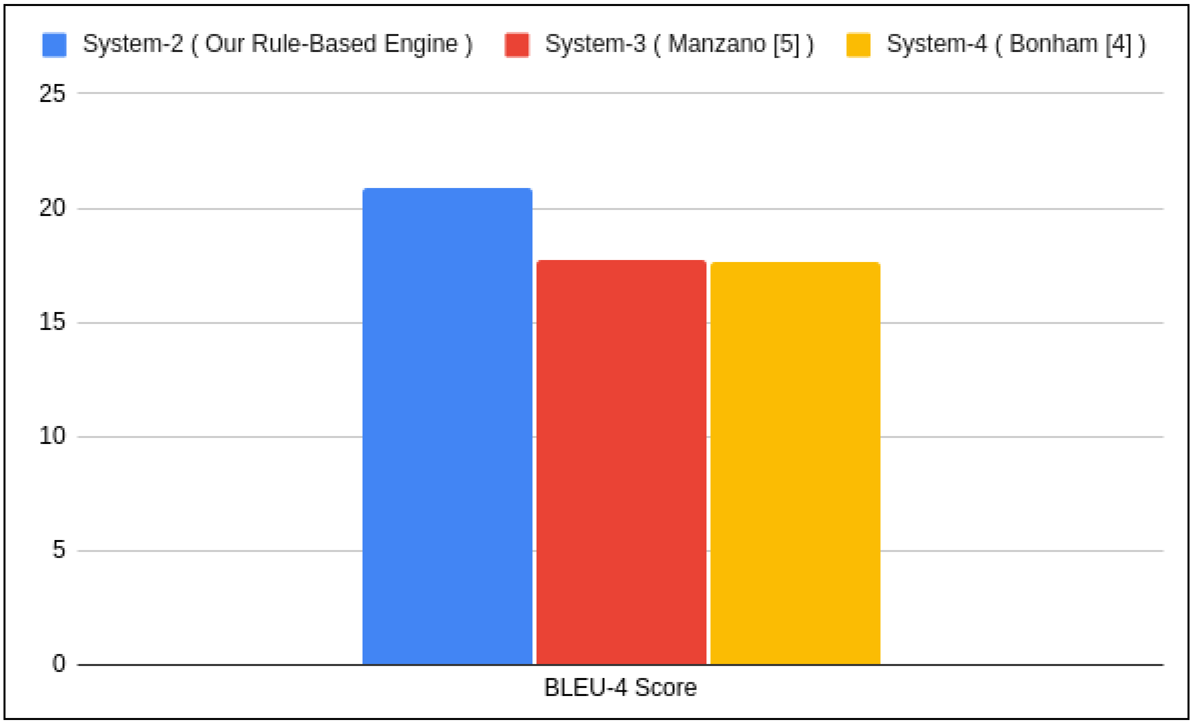

Figure 4.2: Comparison of our rule-based engine with ASL based studies

4.3 Glossing Results

The glossing results of different variations of sentences given in Table 3.2 are given in table below

1st Category	2nd Category	3rd Category	English Sentence	Human Translated	Results Generated By Our Rule-Based Engine
Wh - Questions	Regular Wh - Questions		What kind of soup do you like?	soup, you like what-kind?	soup you like what kind?
	How Wh - Questions		How often do you go to the library?	library, you go how-often?	library you go how often?
	Choice Wh - Questions		Which do you prefer to drink, water, milk, pop, or beer?	water, milk, pop, beer, you prefer drink which?	water milk pop or beer? you prefer drink which .
Polar Questions	Regular Polar Question		Do you have a backpack?	backpack, have you?	backpack have you?
	If Polar Question	Regular If	If your dog gets	suppose your dog	suppose your dog

		Question	sick do you take it to the veterinarian?	sick, you carry vet?	sick, it you take veterinarian?
		Choice If Question	If you go to church, do you wear pants or do you wear a dress?	suppose you go church, pants or dress you which?	suppose you go church, pants you wear or dress you wear ?
	Choice Polar Question		Do you prefer hamburgers or hotdogs?	you prefer hamburger or hotdog?	hamburgers or hotdogs, you prefer?
	And Polar Question		Do you like green eggs and ham?	you like green egg and ham?	you like, green eggs and ham?
Simple Sentences	Regular Sentence		Yesterday I bought a dog.	yesterday dog i buy	yesterday dog i buy.
	Rhetorical Sentence	Regular Rhetorical Sentence	My sister got into a car accident because of black ice.	my sister car accident why? black ice	my sister car accident get why? black ice.
		And Rhetorical Sentence	I love the fall because of the rain and the wind.	fall i love why? rain wind	fall i love why? the rain and the wind
	But Sentence		His wife has red curly hair, but I don't know her name.	his wife red curly hair name i do not know	his wife red have curly hair, her name i not know
	Comma Separated Sentence		I want a new bowl, this one is old.	this bowl it old new bowl i want	new bowl i want, this one old.
	And Sentence	Type 1	The moon and stars are bright tonight.	tonight moon star bright	tonight bright, the moon and stars
		Type 2	The mountain is here and the farm is just below it.	mountain there/ farm there	mountain here, farm just it.
		Type 3	He pays me the money and I go out and buy the food.	he pay-me money. I go-out buy food.	money he pay me, i go out, food buy.

Table 4.5: Comparison of glossing results translated by human and those generated by our rule-based engine

4.4 Discussion

As it can be seen in Figure 4.2 above the BLEU score for our system has outperformed the 4 major systems. The reason for System-1 to have such a high BLEU score is that it is only calculated on data that is related to a specific domain (ID card offices) [8].

The comparison of ASL-based studies in Table 4.4 shows that our rule-based engine has the highest BLEU-4 score.

The cumulative BLEU-4 score of 20.85 that we achieved was calculated using max n-gram = 4, however, upon looking at our results on the individual sentences from test data we found that 8.8 % of sentences had a cumulative and individual BLEU-4 score of 0.00, although many of the machine translations were 100 % accurate (when compared to human translations) for the penalized sentences. The reason that the penalized sentences had a BLEU-4 score of zero was because of their short length. A few examples of penalized sentences with their machine translation are given below

Human Translated	Results Generated By Our Rule-Based Engine	BLEU-4 score
you full you?	you full?	0.00
waiter where?	waiter where?	0.00
student he?	student he?	0.00
open door	door open.	0.00
asl teacher are you?	teacher you?	0.00

Table 4.6: Examples of penalized sentences with n-gram = 4

Considering the fact that a BLEU score with n-gram = 4, would penalize our machine translations, we calculated the results of individual sentences using max n-gram = 3. The cumulative BLEU score we got was 29.76. and the percentage of sentences with a BLEU score of 0.00 reduced to 1.2 %. The penalized sentences with max n-gram = 3 were either totally wrong or too short.

From the above results, we concluded that for ASL glossing, a BLEU score with n-gram = 4 might not be a very good option as can be seen in Table 4.6. ASL glossing generally reduces the number of words from the original English sentence and therefore, considering the results above, BLEU score with n-gram = 3 would be a better option for ASL glossing.

CHAPTER 5: FUTURE WORK

Research in the field of ASL is being done for a number of years but still there is a lot to be done in this field. From the research point of view, the researchers can work on combining a rule-based engine with Artificial intelligence to predict results in a better way.

Field-related ASL rules can also be built that will not only provide more insight into the subject at hand but will give better results for that specific field.

CHAPTER 6: CONCLUSION

The motivation behind this research was to reduce the communication gap between common people and the deaf community. We are quite thankful to Allah Almighty that to the best of our knowledge we were able to **attain the highest BLEU score (20.85) as compared to studies based on ASL**. The research has not only helped me to sharpen my technical skill but it has also helped me to understand the day-to-day struggle of the deaf community. This research will surely pave way for further research opportunities in this field thus leading to the reduction of the communication gap between us and our deaf brothers and sisters.

We also conclude that a better metric system in the field of Sign Language will help the researchers to get better results as the commonly used BLEU score might not be a good metric overall. Moreover, custom metrics can also be made for different sign languages to help future researchers in the respective field.

We also conclude that there is a need for more publications from the ASL interpreters, teachers, and professionals in the field of Text-to-Gloss as finding and extracting the data manually was a cumbersome task and further publications can help future researchers to have easy access to data.

REFERENCES

[1] W. C. Stokoe, 1960, “10th anniversary classics sign language structure: An outline of the visual communication systems of the american deaf”

[2] Tmar, Zouhour I& Othman, Achraf I& Jemni, Mohamed, 2013, “A rule-based approach for building an artificial English-ASL corpus” 1-4.10.1109/ICEESA.2013.6578458.

[3] J. Bungeroth, D. Stein, M. Zahedi, and H. Ney, 2006, “A german sign language corpus of the domain weather report,” in In 5th international, p. 29.

[4] Mary Elizabeth Bonham, 2015, “English to ASL Gloss Machine Translation”

[5] Daniel Moreno Manzano, 2017, “ENGLISH TO ASL TRANSLATOR FOR SPEECH2SIGNS”

[6] Nikolaos Arvanitis, Constantinos Constantinopoulos, Dimitrios Kosmopoulos, 2019, “Translation of Sign Language Glosses to Text Using Sequence-to-Sequence Attention Models”, 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)

[7] Dilek Kayahan and Tunga Gungor, 2019, “A Hybrid Translation System from Turkish Spoken Language to Turkish Sign Language”, IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)

[8] Rub ́en San-Segundo, R Barra, R C ́ordoba, L Fernando D’Haro,F Fern ́andez, Javier Ferreiros, Juan Manuel Lucas, Javier Mac ́ıasGuarasa, Juan Manuel Montero, and Jos ́e Manuel Pardo. Speech to sign language translation system for spanish. Speech Communication, 50(11-12):1009–1020, 2008.

[9] Stoll, S., N. C. Camgöz, S. Hadfield and R. Bowden, 2018, “Sign Language Production using Neural Machine Translation and Generative Adversarial Networks", BMVC.

[10] Kayo Yin, Jesse Read, 2020, “Better Sign Language Translation with STMC-Transformer”

[11] Jiangbin Zheng, Zheng Zhao, Min Chen, Jing Chen, Chong Wu , Yidong Chen , Xiaodong Shi, and Yiqi Tong, 2020, “An Improved Sign Language Translation Model with Explainable Adaptations for Processing Long Sign Sentences”

[12] Jens Forster, Christoph Schmidt, Oscar Koller, Martin Bellgardt, and Hermann Ney, 2014, “Extensions of the Sign Language Recognition and Translation Corpus RWTH-PHOENIX-Weather”. In Proceedings of the Ninth International Conference on

Language Resources and Evaluation (LREC'14), pages 1911–1916, Reykjavik, Iceland. European Language Resources Association (ELRA).

[13] Sang-Ki Ko, Chang Jo Kim, Hyedong Jung, and Choongsang Cho, 2019, "Neural Sign Language Translation based on Human Keypoint Estimation".

[14] B. Woll, R. Sutton-Spence, and D. Waters, 2004, "Echo data set for british sign language (bsl)," in Department of Language and Communication Science,City University (London).

[15] B. Bergman and J. Mesch, 2004, "Echo data set for swedish sign language (ssl)," in Department of Linguistics, University of Stockholm.

[16] O. Crasborn, E. Kooij, A. Nonhebel, and W. Emmerik, 2004, "Echo data set for sign language of the netherlands (ngt)" in Department of Linguistics,Radboud University Nijmegen, 2004.

[17] V. Athitsos, C. Neidle, S. Sclaroff, J. Nash, R. Stefan, A. Thangali, H. Wang, and Q. Yuan, 2010, "Large lexicon project: American sign language video corpus and sign language indexing/retrieval algorithms," in Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages:Corpora and Sign Language Technologies, LREC.

[18] A. Othman, Z. Tmar, and M. Jemni, "Toward developing a very big sign language parallel corpus," Computers Helping People with Special Needs, p. 192–199, 2012.

[19] spaCy models (https://spaCy.io/models/en#en_core_web_sm)

[20] spaCy part-of-speech tagging (https://spaCy.io/usage/linguistic-features#pos-tagging)

[21] Label Scheme for small english model (https://spaCy.io/models/en)

[22] spaCy POS tags list (https://machinelearningknowledge.ai/tutorial-on-spaCy-part-of-speech-pos-tagging/)

[23] ASL Grammar - The Workbook" (2018), by Rochelle Barlow.

[24] BLEU score (https://cloud.google.com/translate/automl/docs/evaluate#bleu)

[25] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to Sequence Learning with Neural Networks. In Advances in neural information processing systems, 2014.

[26] Thang Luong, Hieu Pham, and Christopher D Manning. Effective Approaches to Attention-based Neural Machine Translation. In Conference on Empirical Methods in Natural Language Processing (EMNNLP), 2015.

[27] Sarah Ebling, Necati Cihan Camgoz, Penny Braem, Katja Tissi, Sandra Sidler-Miserez, Stephanie Stoll, Simon Hadfield, Tobias Haug, Richard Bowden, Sandrine Tornay, Marzieh

Razavi, and Mathew Magimai-Doss. Smile swiss german sign language dataset, 02 2018.

[28] Jeffrey Pennington, Richard Socher, and Christopher D Manning. Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pages 1532–1543, 2014.

[29] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. Transactions of the association for computational linguistics, 5:135–146, 2017.

[30] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. Advances in neural information processing systems, 30,2017.